

**The Fermilab Computing Farms in 2000**

Troy Dawson, Jim Fromm, Lisa Giacchetti, Terry Jones
Tanya Levshina, Igor Mandrichenko, Ramon Pasetes
Karen Shepelak, Dane Skow, Steven Timm, Stephen Wolbers

May 15, 2001

Introduction

The year 2000 was a year of evolutionary change for the Fermilab computer farms. Additional compute capacity was acquired by the addition of PCs for the CDF, D0 and CMS farms. This was done in preparation for Run 2 production and for CMS Monte Carlo production. Additional I/O capacity was added for all the farms. This continues the trend to standardize the I/O systems on the SGI O2x00 architecture. Strong authentication was installed on the CDF and D0 farms.

The farms continue to provide large CPU resources for experiments and those users whose calculations benefit from large CPU/low IO resources. The user community will change in 2001 now that the 1999 fixed-target experiments have almost finished processing and Run 2, SDSS, miniBooNE, MINOS, BTeV, and other future experiments and projects will be the major users in the future.

The year in review

The original PC farm (fnpc1-37) continued to provide large computing for the user community during 2000. Those nodes are now at the point where their replacement makes sense, given their somewhat slow speed (333 MHz), small memory (128 MB) and small disk (6 GB). Modern applications will not run well on those machines. All of the approximately 150 500 MHz machines ran well and served the CDF, D0 and non-Run 2 community quite well.

Three acquisitions were made in 2000. 40 dual-Pentium III 750 MHz PCs were purchased for CMS computing, 53 dual-Pentium III 750 MHz PCs were purchased for D0 computing, and 40 dual-Pentium III 800 MHz PCs were purchased for CDF

computing. All machines came with at least 512 MB of memory (1 GB for D0) and 40 GB of user disk. The I/O node fnsfh was upgraded to a SGI O2000. The 5 SGI O200 machines were decommissioned from the farms.

The utilization of the farms during 2000 was much higher than previous years. It was driven by the demand and was not limited by the available compute power of the farms. E871 continued to process their 1999 data, E781 started to completely process the 1996-97 data, and CDF and D0 processed test data, Monte Carlo samples and Commissioning run data.

NuTeV ran simulation jobs during the first half of 2000 and made heavy use of the PC farms. A large sample of events was generated. KTeV generated large simulation samples for the ϵ'/ϵ measurement.

BTeV generated large samples of simulated events during the first half of 2000. These were used to optimize the design of the detector and to study the capability of the proposed detector for B physics. A simulation of Tevatron performance was also performed on the farms. The Sloan Digital Sky Survey began to use the farms to process data (actually reprocess data). This extended the experiment's capabilities and was very useful. CMS ran the first Monte Carlo samples on the shared farms in 2000. Later the CMS calculations moved to the dedicated CMS farms, which are more properly part of the CMS User Facility at Fermilab.

CPU utilization

Table 1 provides the summary of CPU time (in SpecInt95 units) for the whole farm in 2000. A plot of the total CPU utilization (including all previous years of the UNIX farms) is shown in Figure 1.

Table 2 and Figure 2(a-e) show the utilization for each of the many experiments that have used the farms during 2000. E871 was the largest user of the farms, reflecting the huge amount of data that was processed on the farms. KTeV, BTeV, D0, CDF, Beams Division and E781 all used substantial computing on the farms in 2000. NuTeV, CMS, MiniBooNE and SDSS used a fair amount of computing as well. The farm as a whole delivered about 4 times as much total computing (in units of SpecInt95-years) in 2000 than in 1999. This reflects the increase in computing power and an increase in demand for the computing cycles. The total delivered

(1565 SpecInt95-years) is approximately 25% of the total CPU capacity available on the farm at the beginning of the year. CDF and D0 did not use their farms heavily, therefore 2/3 of the farm capacity was mainly used for tests and smaller jobs. The other 1/3 was used much more heavily.

Table 3 is a list of total CPU used (in units of SpecInt95-years) for the past 2 years and for the entire history of the UNIX farms (since the year 1991). The total CPU delivered increased by 67% in just one year, reflecting the huge increase in computing power available. The largest single user of the farms (integrated use since 1991) is now E871 – not a surprise given the huge number of events written to tape by the experiment. D0 is second, E831 third and CDF fourth.

Table 1 – Total CPU use on the Farms – 2000

<u>Month</u>	<u>CPU delivered</u> (SpecInt95-month/month)
January	1007
February	709
March	1354
April	2214
May	2060
June	1511
July	1115
August	855
September	1439
October	1825
November	2598
December	2600

Table 2 – CPU use by experiment – 2000

<u>Experiment</u>	<u>CPU time</u> (SpecInt95-yrs)
E871	413
KTeV	234
BTeV	205
D0	194
CDF	142
Beams	95
E781	95
NuTeV	86
CMS	49
MiniBooNE	32
SDSS	15
Theory	4
NUMI	1
TOTAL	1565

Table 3
Integrated Farm Use
(In units of SpecInt95-years)
Through December, 2000

Experiment	1999 (1991-2000)	2000Total(1991-2000)	
E871	209	413	718
D0 offline		194	490
E831	29		419
CDF		142	303
KTeV		234	234
E781		95	212
E706			209
BTeV		205	205
E791			176
Auger	52		171
NuTeV	80	86	166
Beams	16	95	125
D0 MC			62
E665			60
Theory	36	4	59
E771			54
CMS		49	49
MiniBooNE		32	32
E866			30
NUMI	2	1	28
E789			25
E687			25
Minos			14
Recycler			13
E872	2		7
E760			3
E835			3
E731			2
Magnet			2
Total	425	1565	3911

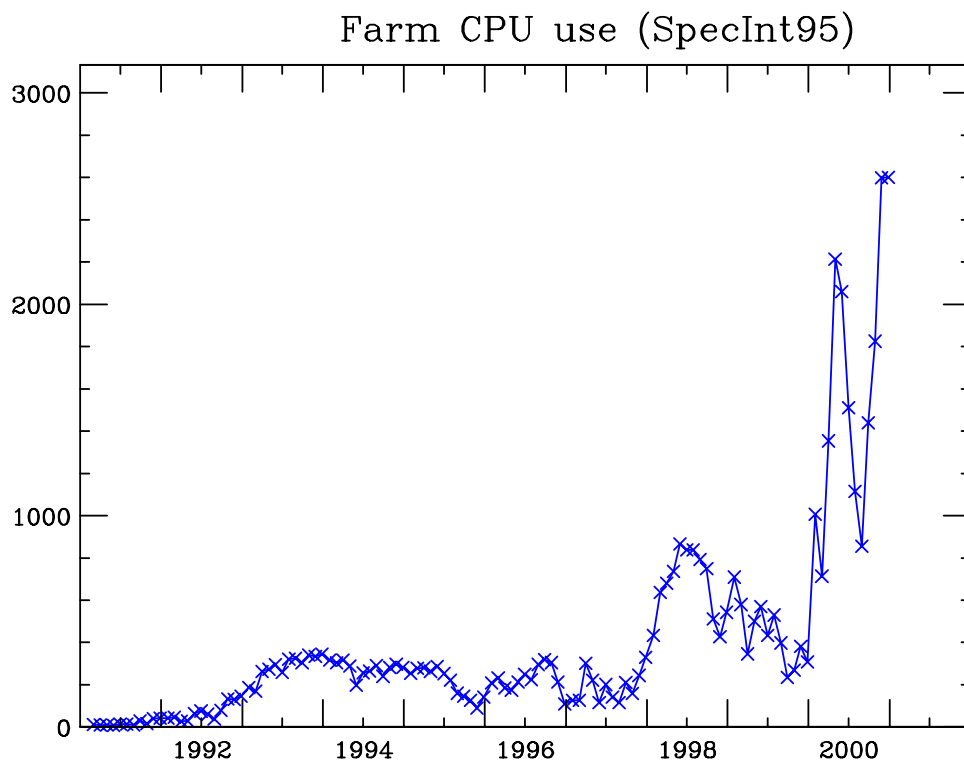


Figure 1.

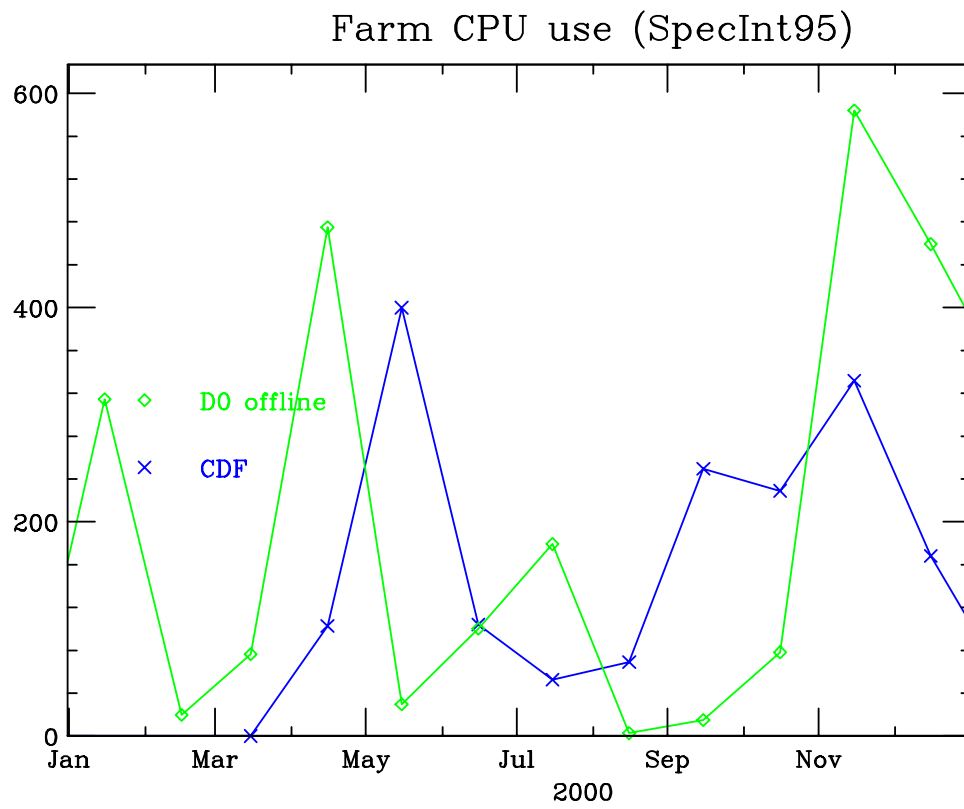


Figure 2(a).

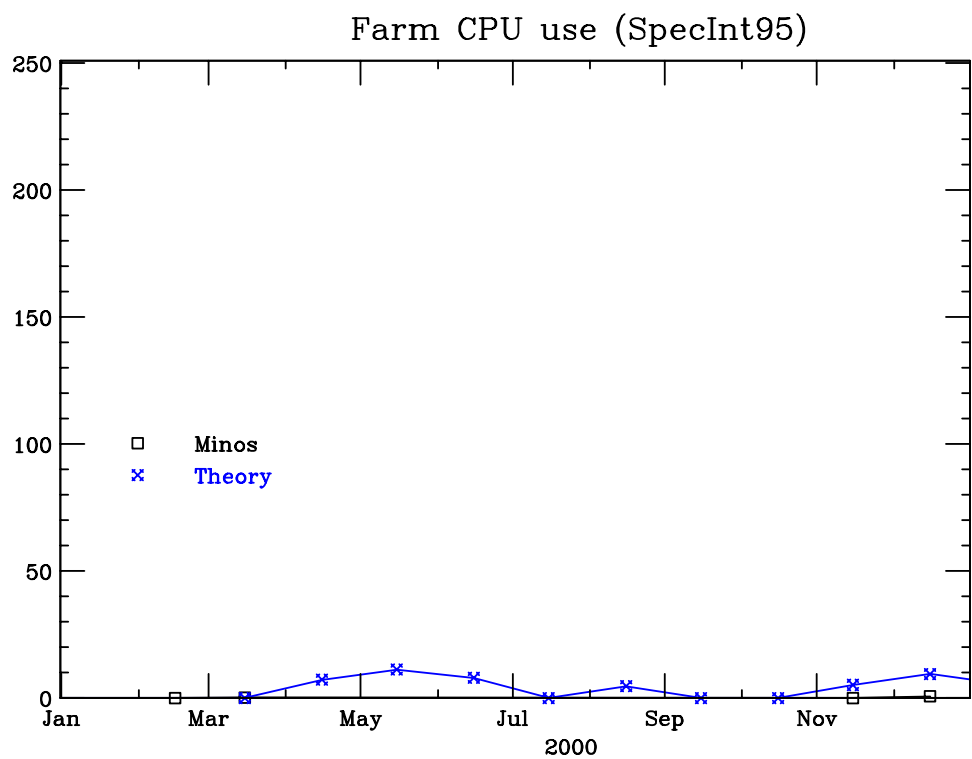


Figure 2(b).

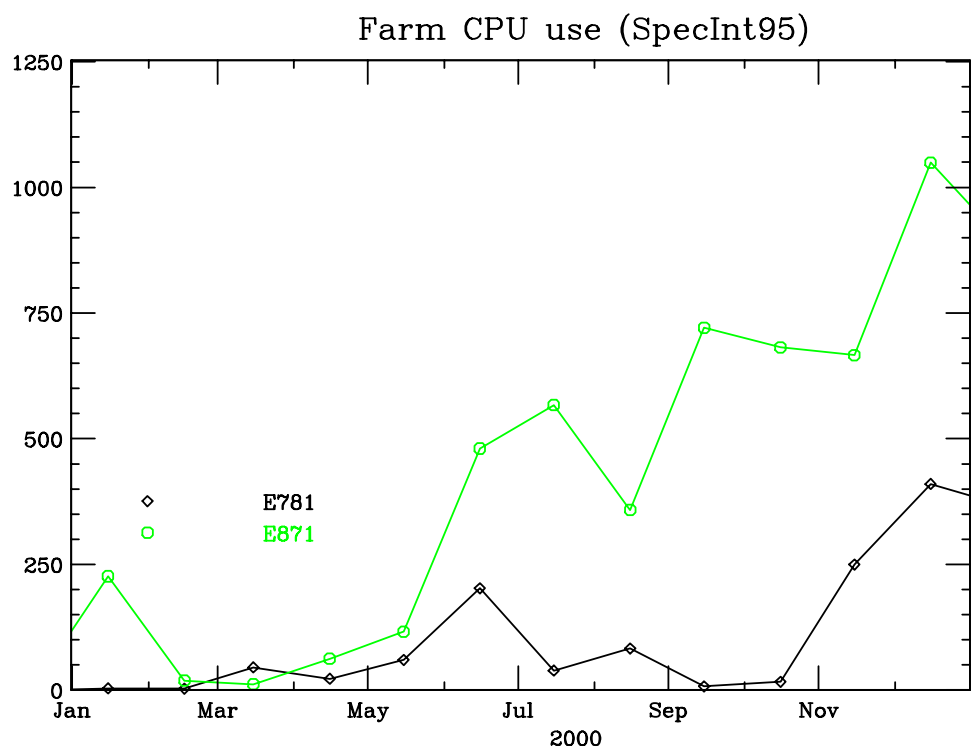


Figure 2(c).

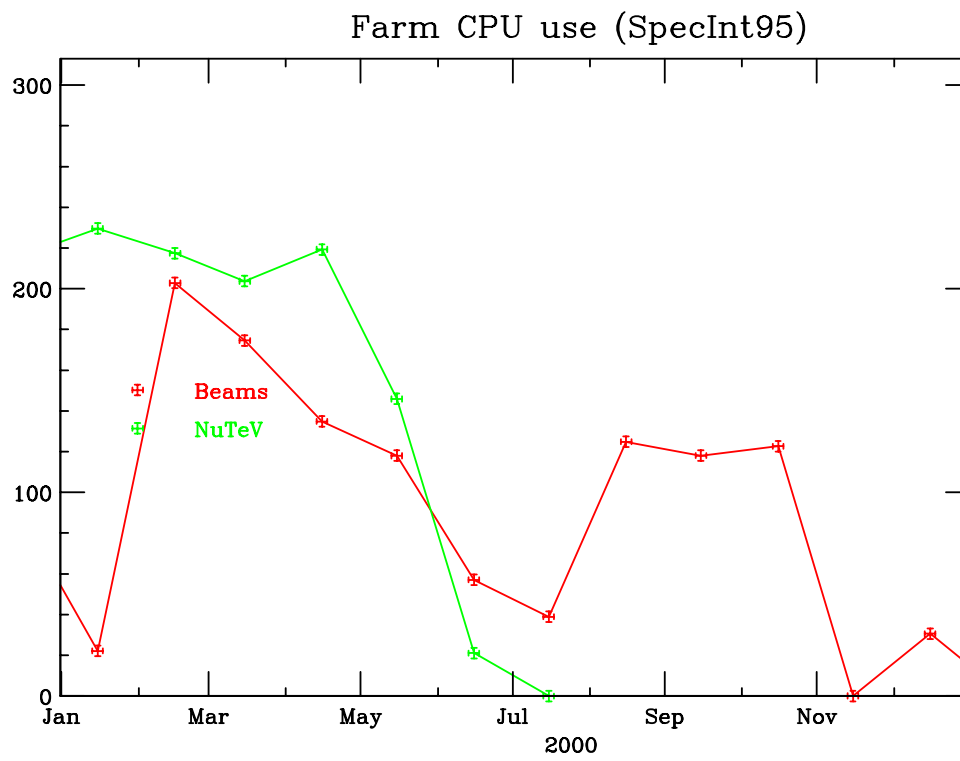


Figure 2(d).

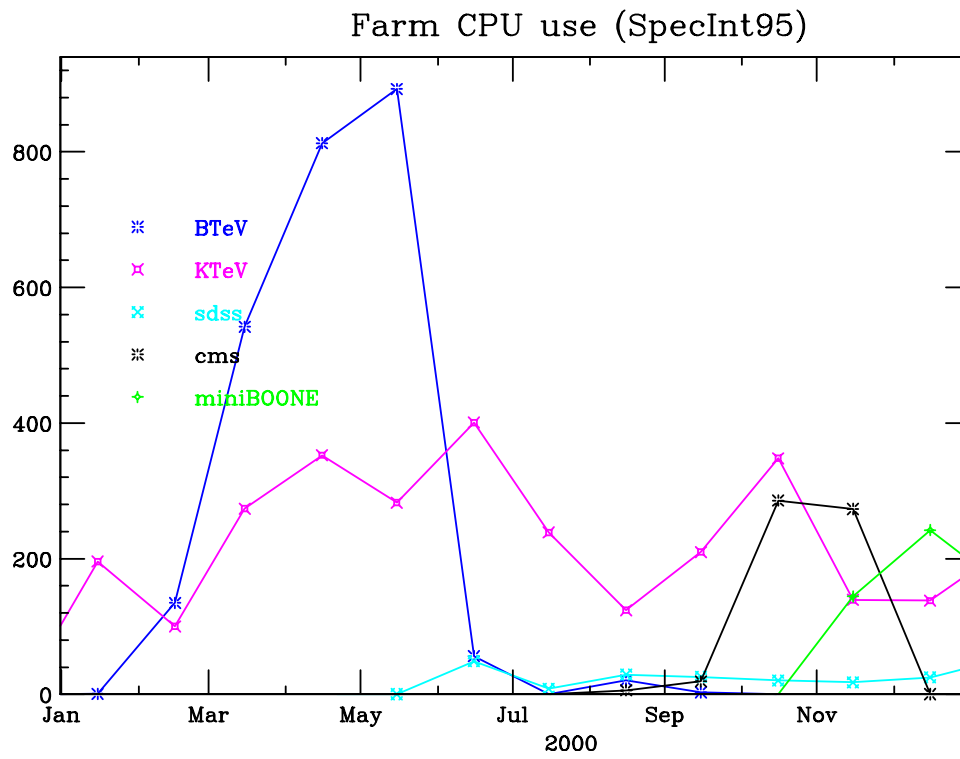


Figure 2(e).

Allocations

The farms are allocated more dynamically than in the past. Instead of dedicating specific nodes to individual experiments or groups FBSNG (the farms batch scheduler (<http://www.isd.fnal.gov/fbsng>)) allows a much more flexible and dynamic allocation of resources compared to FBS and to static allocations which have been used in the past. There are constraints and priorities for the resources and these are tuned to maximize the utilization while at the same time ensuring that the highest priority tasks get the resources that they need.

Therefore, the allocation for any experiment or user is changing within the limits allowed. Table 4 gives the total capacity of different types of PCs on the 4 distinct farms – CDF, D0, CMS and “fixed-target farms”. The total available CPU power is 16,575 SpecInt95, much greater than the total available a year ago (6,400). Table 5 lists the SpecInt95 ratings for the machines used at Fermilab.

Table 4
Current Farm Capacity, number of nodes and total CPU
All Nodes are duals

Farm	PC(333)	PC(500)	PC(750)	PC(800)	SpecInt95
CDF		48		40	4998
D0		47	53		5702
CMS			40		2856
FT	36	50			3019
TOTALS	36	145	93	40	16575

Table 5—SpecInt95 Rating for PCs

PC(333)	13.6
PC(500)	20.4
PC(750)	35.7
PC(800)	38.0

Plans

2001 will see the ramp up of the use of all the systems currently in production. This will be driven by Run 2, which will begin to take a significant amount of data by mid-2001. CMS will also use systems heavily as part of the Monte Carlo production. E871 and E781 will complete the reconstruction of their data by mid-2001. KTeV will continue to use systems to create Monte Carlo samples. BTeV expects to use a significant amount of CPU time to further improve understanding of the proposed detector and to complete the TDR (Technical Design Report). MiniBooNE will likely continue to use the farms for simulation jobs. SDSS will continue to use the farm for processing of data.

The CDF farms will increase in 2001 with a purchase of 64 PCs of at least 1 GHz speed. A purchase near the end of the calendar year for more PCs is possible, though it depends on the ramp-up of data-taking and the time required for event reconstruction. The D0 farms will increase as well in 2001 with the timing dependent on data-taking ramp-up. The “Fixed-Target farms” will grow with the addition of 40 PCs in summer 2001. The oldest PC farm (333 MHz machines) will likely be retired in 2001 from active farm use and will move to other uses.

A new switch will be acquired for the “Fixed-Target farms” to replace the old Catalyst 5500. A new switch for the D0 farms is a good possibility. A new switch for CMS computing is also likely in 2001.

Finally, the long-awaited move to robotic storage for essentially all data access on the farms will occur in late 2001. The last experiments to use the farm for manual tapemounts (E781, E871 and E835) will finish processing by mid-2001. All new experiments (CDF, D0, BTeV, MiniBooNE, MINOS, CMS) will likely never use an operator tapemount. This will allow the I/O systems for the farms to change as the tapedrives can be slowly removed.